Systèmes de réaction-diffusion et HPC: application aux AVC

Thierry Dumont[1], Violaine Louvet[1]
Stéphane Descombes [2], Max Duarte [3], Marc Massot[4]

[1]Institut Camille Jordan/CNRS - Université Claude Bernard Lyon 1
[2]J. Dieudonné Laboratory, Nice-Sophia Antipolis University - France
[3]Center for Turbulence Research - Stanford University - USA
[4]EM2C Laboratory - Ecole Centrale Paris - France

Rouen, 19 juin 2013

1. Contexte et enjeux sociétaux
2. Modélisation de l'AVC
3. Stratégies numériques
4. Détails de la mise en œuvre
5. Problématique de la parallélisation
6. Conclusions et perspectives

Sommaire

1. Contexte et enjeux sociétaux
2. Modélisation de l'AVC
3. Stratégies numériques
4. Détails de la mise en œuvre
5. Problématique de la parallélisation
6. Conclusions et perspectives

Collaboration inter-disciplinaires

- Maths
  - T. Dumont (IR UCBL : calcul scientifique).
  - E. Grenier (Pr. ENSL : modélisation, maths)
  - S. Descombes (Pr. Nice : analyse numérique)
  - G. Chapuisat (Mc. Marseille : modélisation, maths).
  - M. Massot (Pr. ECP : analyse numérique),
- Médecine
  - J-P. Boissel (Pr. Pharmacologue, Lyon)
  - M-A. Dronne (Mc. Pharmacologue, Lyon)
  - beaucoup d'autres personnes ...

Collaboration inter-disciplinaires

- Maths
  - T. Dumont (IR UCBL : calcul scientifique).
  - E. Grenier (Pr. ENSL : modélisation, maths)
  - S. Descombes (Pr. Nice : analyse numérique)
  - G. Chapuisat (Mc. Marseille : modélisation, maths).
  - M. Massot (Pr. ECP : analyse numérique),
- Médecine
  - J-P. Boissel (Pr. Pharmacologue, Lyon)
  - M-A. Dronne (Mc. Pharmacologue, Lyon)
  - beaucoup d'autres personnes ...

Contexte

- Projet INRIA Numed (`http://www.umpa.ens-lyon.fr/~numed/`).
- Projet ANR Séchelles (`http://math.unice.fr/~sdescomb/Sechelles.html`).
- **I**nstitut de **M**édecine **TH**éorique (`http://imth.univ-lyon1.fr`).
- Projet ANR **AVC** **I**n **SI**lico.

Stroke

- AVC ischémiques : résulte de l'occlusion d'un ou plusieurs vaisseaux sanguins dans le cerveau.
- Un problème de santé publique majeur car ils représentent la deuxième cause de mort au niveau mondial et la première cause de handicaps acquis chez l'adulte.

Stroke

- AVC ischémiques : résulte de l'occlusion d'un ou plusieurs vaisseaux sanguins dans le cerveau.
- Un problème de santé publique majeur car ils représentent la deuxième cause de mort au niveau mondial et la première cause de handicaps acquis chez l'adulte.

Enjeux sociétaux

- La compréhension des mécanismes de l'AVC est essentielle pour développer des stratégies thérapeutiques.
- Neuro-protecteurs (qui bloquent la cascade ischémique) : convaincants chez le rat, désastreux chez l'homme. Pourquoi ?
- Etre capable de prédire les dommages et d'adapter la thérapie en temps réel.

Besoin d'un modèle et de simulations

Sommaire

1 Contexte et enjeux sociétaux

2 Modélisation de l'AVC

3 Stratégies numériques

4 Détails de la mise en œuvre

5 Problématique de la parallélisation

6 Conclusions et perspectives

Composition du cerveau

Les tissus cérébraux sont composés de 2 types de cellules :

- neurones,
- cellules gliales,

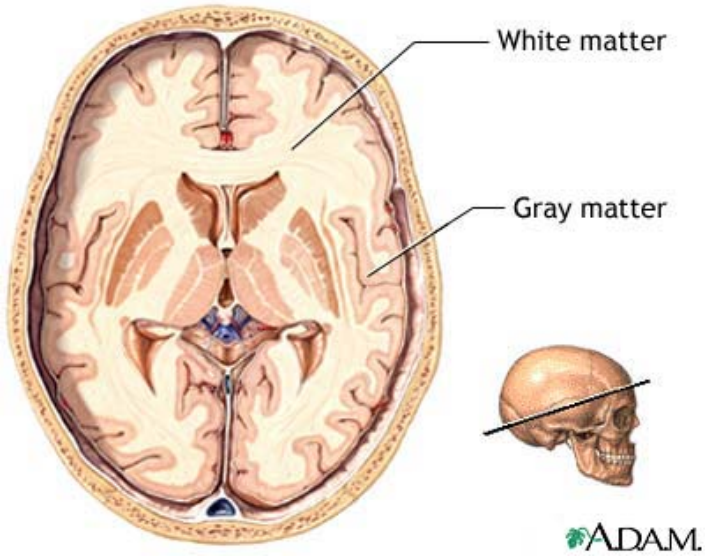
et de milieu intracellulaire.

Composition du cerveau

Les tissus cérébraux sont composés de 2 types de cellules :

- neurones,
- cellules gliales,

et de milieu intracellulaire.



2 domaines qui diffèrent par leur composition en cellule gliales et neuronales :

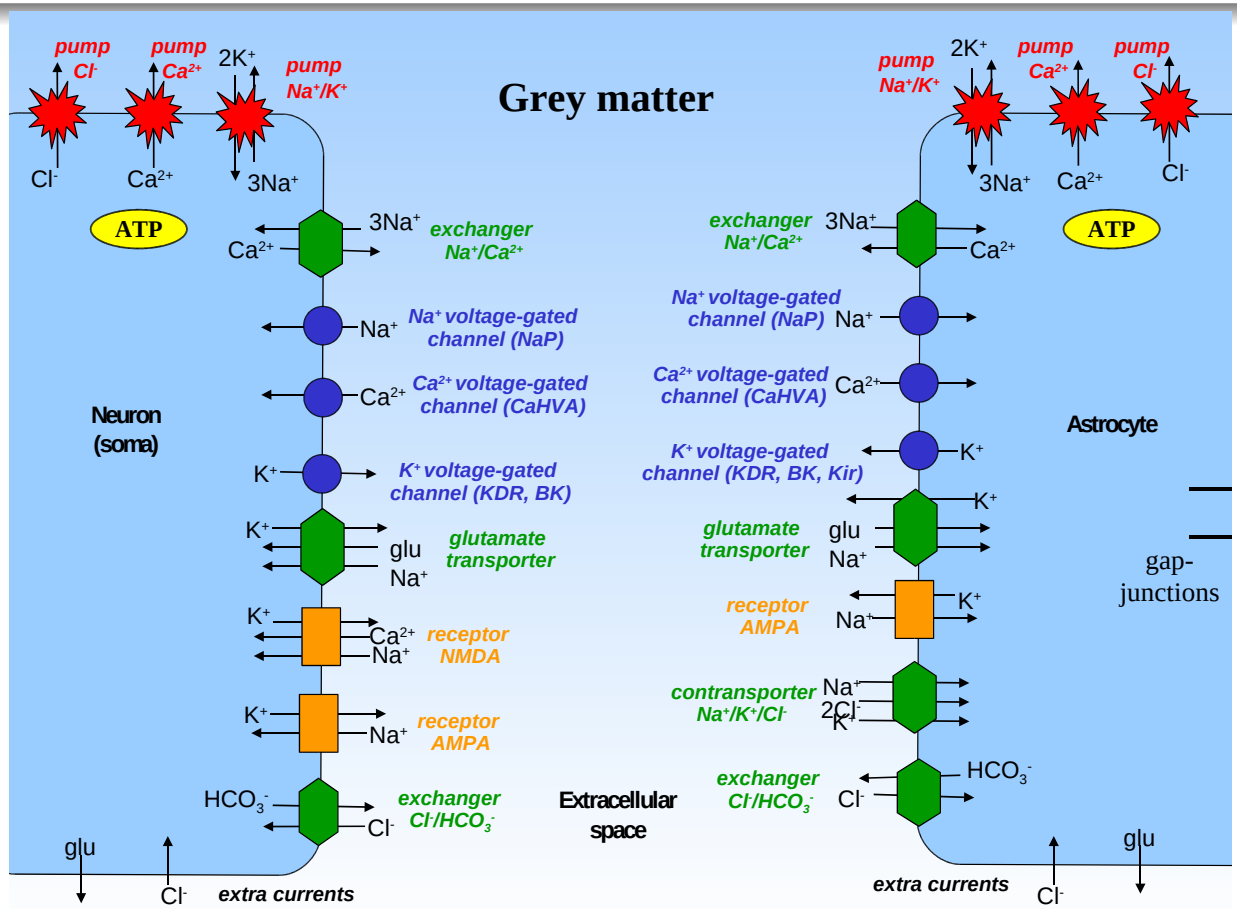
- matière blanche : astrocytes et sommas neuronaux
- matière grise : oligodendrocytes et axones neuronaux.

Chez le rat : pas de matière blanche et une géométrie très différente

Modélisation de l'AVC (M-A. Dronne, E. Grenier, J-P. Boissel)

Modèle électrophysiologique

Les échanges ioniques sont le mécanisme principal menant à la mort cellulaire.



Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire
- augmentation de la concentration en Ca^{2+} => activation enzymatique => nécrose

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire
- augmentation de la concentration en Ca^{2+} => activation enzymatique => nécrose
- augmentation du glutamate dans l'espace extracellulaire (toxique)

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire
- augmentation de la concentration en Ca^{2+} => activation enzymatique => nécrose
- augmentation du glutamate dans l'espace extracellulaire (toxique)
- augmentation de la concentration en K^{+} qui diffuse dans le milieu extracellulaire et de Ca^{2+} dans les astrocytes => création d'ondes de dépression corticale => propagation des dégâts loin de la zone ischémiée

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire
- augmentation de la concentration en Ca^{2+} => activation enzymatique => nécrose
- augmentation du glutamate dans l'espace extracellulaire (toxique)
- augmentation de la concentration en K^{+} qui diffuse dans le milieu extracellulaire et de Ca^{2+} dans les astrocytes => création d'ondes de dépression corticale => propagation des dégâts loin de la zone ischémiée

Modèle physiologique

- Occlusion d'un vaisseau sanguin => moins d'énergie => panne des pompes ioniques
- altération du potentiel membranaire => ouverture des canaux ioniques dépendant du voltage (portes)
- augmentation du volume cellulaire
- augmentation de la concentration en Ca^{2+} => activation enzymatique => nécrose
- augmentation du glutamate dans l'espace extracellulaire (toxique)
- augmentation de la concentration en K^{+} qui diffuse dans le milieu extracellulaire et de Ca^{2+} dans les astrocytes => création d'ondes de dépression corticale => propagation des dégâts loin de la zone ischémiée

Les neuro-protecteurs tentent de piloter les canaux dépendant du voltage pour empêcher la propagation des ondes de dépression corticale.

Système de réaction–diffusion

$$\begin{cases} \dfrac{\partial u_i}{\partial t}(x,t) - \mathrm{div}(\varepsilon_i(x)\mathrm{grad}\ u_i(x,t)) & = f_i(u_1(x,t),\ldots,u_m(x,t)), \\ & 1 \leq i \leq m,\ x \in \Omega, \\ u_i(x,0) = u_i^0(x), & 1 \leq i \leq m,\ x \in \Omega. \end{cases}$$

Système de réaction–diffusion

$$\begin{cases} \dfrac{\partial u_i}{\partial t}(x,t) - \mathrm{div}(\varepsilon_i(x)\mathrm{grad}\ u_i(x,t)) & = f_i(u_1(x,t),\ldots,u_m(x,t)), \\ & 1 \le i \le m,\ x \in \Omega, \\ u_i(x,0) = u_i^0(x), & 1 \le i \le m,\ x \in \Omega. \end{cases}$$

- 3 milieux : neurones, astrocytes et milieu extra cellulaire
- Concentrations ioniques : K^+, Na^+, Cl^-, Ca^{2+} et Glu^- dans chaque milieu (15 unknowns).
- S_n, S_a proportion volumique des cellules. La seule grandeur observable à l'IRM est le volume extra-cellulaire : $1 - S_n - S_a$.
- V_n et V_a : potentiels.

Système de réaction–diffusion

$$\begin{cases} \dfrac{\partial u_i}{\partial t}(x,t) - \mathrm{div}(\varepsilon_i(x)\mathrm{grad}\ u_i(x,t)) & = f_i(u_1(x,t),\ldots,u_m(x,t)), \\ & 1 \le i \le m,\ x \in \Omega, \\ u_i(x,0) = u_i^0(x), & 1 \le i \le m,\ x \in \Omega. \end{cases}$$

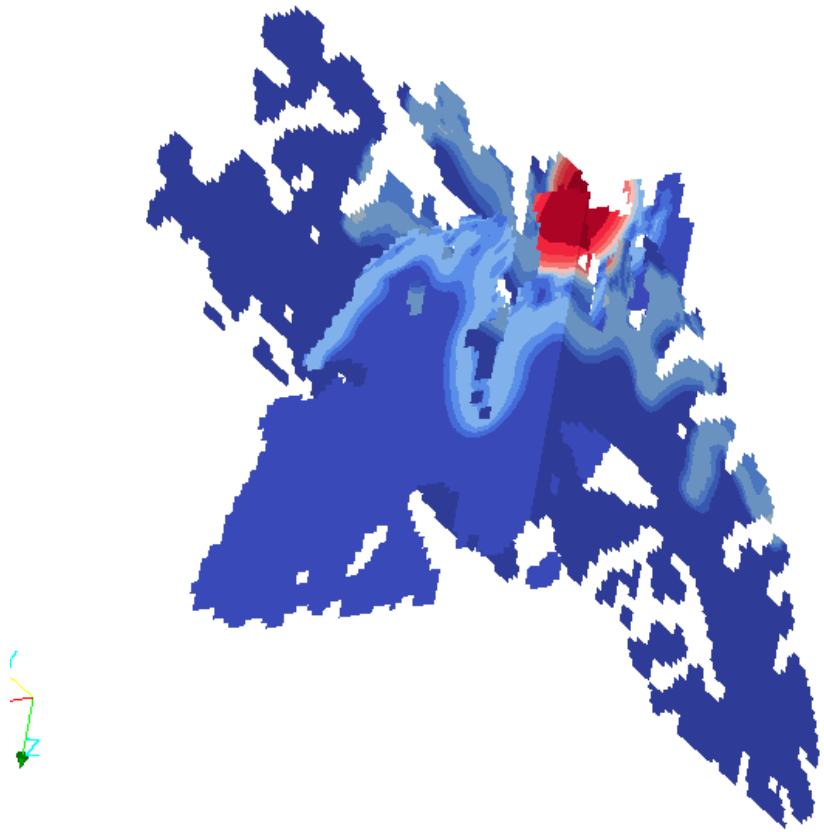
- 3 milieux : neurones, astrocytes et milieu extra cellulaire
- Concentrations ioniques : K^+, Na^+, Cl^-, Ca^{2+} et Glu^- dans chaque milieu (15 unknowns).
- S_n, S_a proportion volumique des cellules. La seule grandeur observable à l'IRM est le volume extra-cellulaire : $1 - S_n - S_a$.
- V_n et V_a : potentiels.

Pas de diffusion pour S_n, S_a, V_n et V_a.
Au total 19 inconnues.

Difficultés numériques

Problème de *grande taille*

- Domaine 3D, complexe => 2.10^7 nœuds au minimum.
- 19 inconnues => $3,8.10^8$ inconnues discrètes.



Difficultés numériques

Problème de *grande taille*

- Domaine 3D, complexe $=>$ 2.10^7 nœuds au minimum.
- 19 inconnues $=>$ $3,8.10^8$ inconnues discrètes.

Difficultés numériques

- Raideur du terme réactif $F = (f_1, f_2, ..., f_m)^t$
  - Echelles de temps très différentes,
  - Spectre du jacobien de F très large : valeurs propres de DF dans l'intervalle $[-10^8, -\varepsilon]$.

  Problème multi-échelles.
- Diffusion faible.
- Front : raideur spatiale.

Sommaire

1 Contexte et enjeux sociétaux

2 Modélisation de l'AVC

3 Stratégies numériques

4 Détails de la mise en œuvre

5 Problématique de la parallélisation

6 Conclusions et perspectives

Méthodes numériques adaptées

$\hookrightarrow$ Besoin de solveurs dédiés aptes à résoudre toutes les échelles de temps et d'espace dans une configuration multi-dimensionnelle.

Méthodes numériques adaptées

$\hookrightarrow$ Besoin de solveurs dédiés aptes à résoudre toutes les échelles de temps et d'espace dans une configuration multi-dimensionnelle.

$\hookrightarrow$ Besoin d' algorithmes adaptés et efficaces sur les architectures de calcul actuelles (et futures).

- Beaucoup de cœurs/threads :
  - Décroissance de la fréquence d'horloge due à la consommation électrique et à la dissipation calorifique
  - Peu de mémoire par cœur

Méthodes numériques adaptées

$\hookrightarrow$ Besoin de solveurs dédiés aptes à résoudre toutes les échelles de temps et d'espace dans une configuration multi-dimensionnelle.

$\hookrightarrow$ Besoin d' algorithmes adaptés et efficaces sur les architectures de calcul actuelles (et futures).

- Beaucoup de cœurs/threads :
  - Décroissance de la fréquence d'horloge due à la consommation électrique et à la dissipation calorifique
  - Peu de mémoire par cœur

Ingrédients numériques

1. Splitting d'operateur.
2. Intégration en temps pour la diffusion.
3. Intégration en temps pour la réaction.
4. Multirésolution.

Splitting d'opérateurs

Séparer la diffusion et la réaction

Splitting d'opérateurs

Séparer la diffusion et la réaction

$$U(t) = T^t U_0 \quad \begin{cases} \partial_t U - \Delta U = \Omega(U) \\ U(0) = U_0 \end{cases}$$

Splitting d'opérateurs

Séparer la diffusion et la réaction

$$U(t) = T^t U_0 \quad \begin{cases} \partial_t U - \Delta U = \Omega(U) \\ U(0) = U_0 \end{cases}$$

Scindé en 2 blocs élémentaires :

$$V(t) = X^t V_0 \quad \begin{cases} \partial_t V - \Delta V = 0 \\ V(0) = V_0 \end{cases} \qquad W(t) = Y^t W_0 \quad \begin{cases} \partial_t W = \Omega(W) \\ W(0) = W_0 \end{cases}$$

Splitting d'opérateurs

Séparer la diffusion et la réaction

$$U(t) = T^t U_0 \quad \begin{cases} \partial_t U - \Delta U = \Omega(U) \\ U(0) = U_0 \end{cases}$$

Scindé en 2 blocs élémentaires :

$$V(t) = X^t V_0 \quad \begin{cases} \partial_t V - \Delta V = 0 \\ V(0) = V_0 \end{cases} \qquad W(t) = Y^t W_0 \quad \begin{cases} \partial_t W = \Omega(W) \\ W(0) = W_0 \end{cases}$$

Idée principale

Découpler les erreurs d'intégration en temps en choisissant des méthodes d'ordre élevé pour chacun des deux sous-problèmes de façon à ce que l'erreur globale soit principalement pilotée par le pas de temps du splitting.

Directions alternées

Formalisme de Lie, méthode du 1er ordre

$$L_1^t\, U_0 = Y^t\, X^t\, U_0$$

$$L_2^t\, U_0 = X^t\, Y^t\, U_0$$

Directions alternées

Formalisme de Lie, méthode du 1er ordre

$$L_1^t\, U_0 = Y^t\, X^t\, U_0$$

$$L_2^t\, U_0 = X^t\, Y^t\, U_0$$

Formalisme de Strang, méthode du 2nd ordre

$$S_1^t\, U_0 = Y^{t/2}\, X^t\, Y^{t/2}\, U_0$$

$$S_2^t\, U_0 = X^{t/2}\, Y^t\, X^{t/2}\, U_0$$

Directions alternées et raideur

- Réduction d'ordre à cause des échelles rapides (voir référence [2])
- Réduction d'ordre dûe aux gradients en espace très élevés (voir référence [5])

Directions alternées et raideur

- Réduction d'ordre à cause des échelles rapides (voir référence [2])
- Réduction d'ordre dûe aux gradients en espace très élevés (voir référence [5])

Comportement de l'erreur, échelles rapides

Erreur locale pour les formalismes de Lie et de Strang :

$$\|E_L(t)U_0\|_2 \leq \left(\frac{C_{L0}\,t^2}{2} + \frac{C_{L1}\,t\sqrt{t}}{3\sqrt{2e}}\right)\|U_0\|_2$$

$$\|E_S(t)U_0\|_2 \leq \frac{(C_{S0}+2C_{S1})\,t^3}{12} + \frac{C_{S2}\,t^2\sqrt{t}}{15\sqrt{2e}} + \frac{C_{S3}\alpha\,t\sqrt{t}}{4}$$

Directions alternées et raideur

Comportement de l'erreur, gradients spatiaux

Dans le cas de gradients en espace élevé, pour le formalisme de Lie, $\exists$ une constante explicite $\theta > 0$ dépendante de $\|U_0\|_{H^1}$ telle que pour $t \leq \theta$, $\|E_L(t)U_0\|_2$ se comporte en t^2 et pour $t \geq \theta$, $\|E_L(t)U_0\|_2$ se comporte en $t\sqrt{t}$.

Directions alternées et raideur

Comportement de l'erreur, gradients spatiaux

Dans le cas de gradients en espace élevé, pour le formalisme de Lie, $\exists$ une constante explicite $\theta > 0$ dépendante de $\|U_0\|_{H^1}$ telle que pour $t \leq \theta$, $\|E_L(t)U_0\|_2$ se comporte en t^2 et pour $t \geq \theta$, $\|E_L(t)U_0\|_2$ se comporte en $t\sqrt{t}$.

On ne peut utiliser que Strang $Y \circ X \circ Y$! Il faut finir par l'opérateur le plus raide

Briques de base : diffusion

Rock4 [1]

Ce qu'on a appris quand on était petit : *« toujours résoudre l'équation de la chaleur avec une méthode implicite »...*

Briques de base : diffusion

Rock4 [1]

Ce qu'on a appris quand on était petit : *« toujours résoudre l'équation de la chaleur avec une méthode implicite »...*

...n'est *pas toujours vrai.*

Briques de base : diffusion

Rock4 [1]

Ce qu'on a appris quand on était petit : *« toujours résoudre l'équation de la chaleur avec une méthode implicite »...*

...n'est *pas toujours vrai.*

Méthode de Runge-Kutta explicite stabilisée avec un domaine de stabilité le long de l'axe réel négatif.
Méthode d'ordre élevé (ordre 4).

Briques de base : diffusion

Rock4 [1]

Ce qu'on a appris quand on était petit : *« toujours résoudre l'équation de la chaleur avec une méthode implicite »...*

...n'est *pas toujours vrai.*

Méthode de Runge-Kutta explicite stabilisée avec un domaine de stabilité le long de l'axe réel négatif.
Méthode d'ordre élevé (ordre 4).

- Dans notre cas : les coefficients de diffusion ε_i sont petits
- Phénomène de diffusion = propagation des perturbations autour des équilibres partiels résultants des schémas réactifs.
- Méthode appropriée à nos pbs de diffusion à cause de la prédominance de valeurs propres réelles négatives de taille raisonnable pour lesquelles la méthode est stable.

Briques de base : diffusion

Intérêts computationnels

- Méthode explicite : pas de résolution de système linéaire, uniquement des produits matrices-vecteurs.
- Le coût est lié directement à la quantité de produits matrice-vecteur, et donc au nombre d'étapes s à chaque pas de temps de la méthode.
- Caractère auto-similaire qui implique un nombre d'étape pratiquement constant au cours de l'évolution.
- Les besoins en mémoire sont directement liés à la taille du maillage.

Briques de base : Réaction

Radau5 [9]

Système raide =>

- Utilisation d'une méthode implicite
- Méthode A–stable et L–stable
- Méthode d'ordre élevé (formellement d'ordre 5, qui peut être réduit à 3 dans le pire des cas) : l'erreur de l'intégration en temps est bornée par celle de la méthode de splitting
- Schéma de Runge-Kutta implicite : résolution de systèmes non linéaires au cours du processus d'intégration : pb du coût calcul et mémoire. Dans notre cas, résolution de systèmes locaux de petite taille.

Briques de base : Réaction

Radau5 [9]

Système raide =>

- Utilisation d'une méthode implicite
- Méthode A–stable et L–stable
- Méthode d'ordre élevé (formellement d'ordre 5, qui peut être réduit à 3 dans le pire des cas) : l'erreur de l'intégration en temps est bornée par celle de la méthode de splitting
- Schéma de Runge-Kutta implicite : résolution de systèmes non linéaires au cours du processus d'intégration : pb du coût calcul et mémoire. Dans notre cas, résolution de systèmes locaux de petite taille.

L'analyse sur l'erreur théorique du schéma de splitting reste valide : les échelles de temps de la réaction et de la diffusion sont correctement résolues au regard des erreurs de l'intégration en temps du splitting.

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.
- Strang S_1^t : localement d'ordre 3.

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.
- Strang S_1^t : localement d'ordre 3.
- Fabriquer une méthode d'ordre localement 2

$$S_{1,\delta}^t = Y^{(1/2-\delta)t} \circ X^t \circ Y^{(1/2+\delta)t}$$

et mesurer la différence entre les deux méthodes.

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.
- Strang S_1^t : localement d'ordre 3.
- Fabriquer une méthode d'ordre localement 2

$$S_{1,\delta}^t = Y^{(1/2-\delta)t} \circ X^t \circ Y^{(1/2+\delta)t}$$

et mesurer la différence entre les deux méthodes.

$$S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \sim \mathcal{O}(\Delta t^2)$$

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.
- Strang S_1^t : localement d'ordre 3.
- Fabriquer une méthode d'ordre localement 2

$$S_{1,\delta}^t = Y^{(1/2-\delta)t} \circ X^t \circ Y^{(1/2+\delta)t}$$

et mesurer la différence entre les deux méthodes.

$$S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \sim \mathcal{O}(\Delta t^2)$$

Pour une tolérance η donné, on cherche à vérifier :

$$\| S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \| < \eta$$

Adaptation du pas de temps de splitting

- Basé sur une estimation de l'erreur locale obtenue en utilisant deux schémas d'ordres différents.
- Strang S_1^t : localement d'ordre 3.
- Fabriquer une méthode d'ordre localement 2

$$S_{1,\delta}^t = Y^{(1/2-\delta)t} \circ X^t \circ Y^{(1/2+\delta)t}$$

et mesurer la différence entre les deux méthodes.

$$S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \sim \mathcal{O}(\Delta t^2)$$

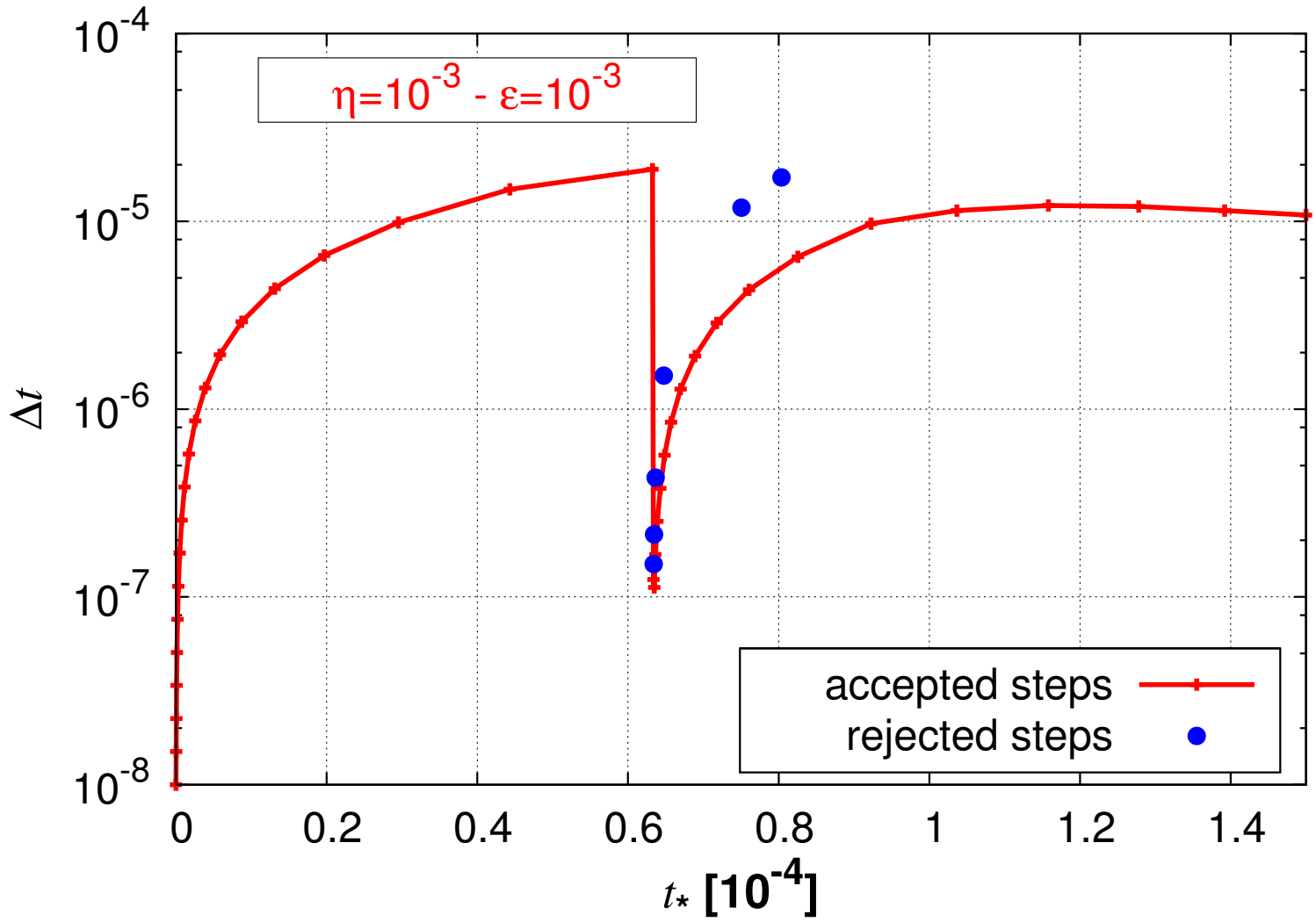
Pour une tolérance η donné, on cherche à vérifier :

$$\| S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \| < \eta$$

Si ce n'est pas le cas, on recalcule Δt :

$$\Delta t_{new} = \Delta t \sqrt{\frac{\eta}{\| S_2^{\Delta t} u_0 - S_{2,\delta}^{\Delta t} u_0 \|}}$$

Adaptation du pas de temps de splitting



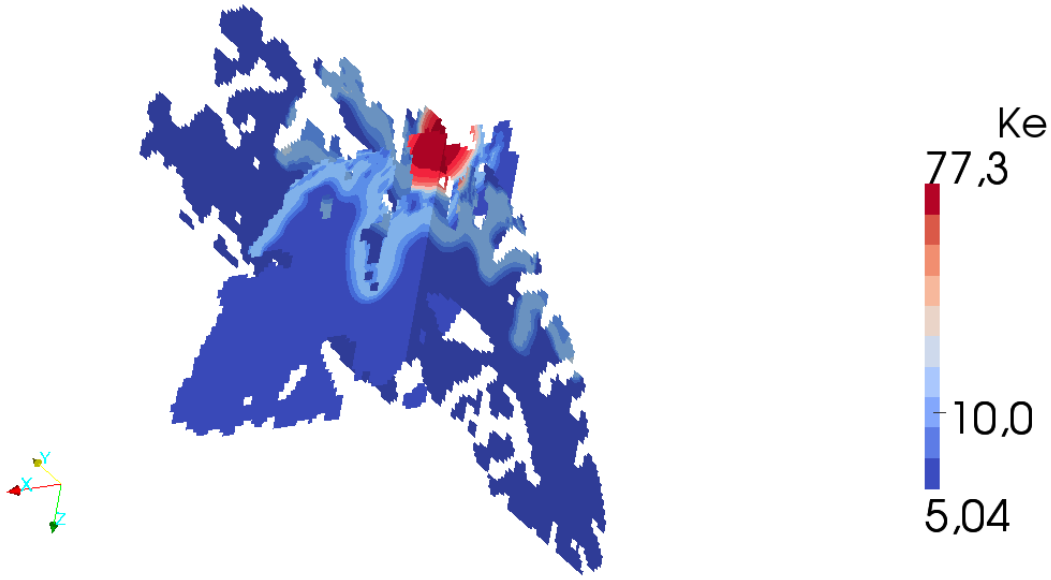
Allumage d'une flamme

Echelles spatiales

Effet local et gradients importants

L'AVC est localisé dans une petite partie du cerveau (au moins au début) et présente des ondes progressives et de forts gradients en espace.

Film



Echelles spatiales

Effet local et gradients importants

L'AVC est localisé dans une petite partie du cerveau (au moins au début) et présente des ondes progressives et de forts gradients en espace.

Sait-on résoudre les échelles spatiales ?

Expérimentation numérique : en 1D, résolution du système de l'AVC avec un solveur d'EDO d'ordre élevé => solution « exacte » du système discrétisé en espace

Au moins 1000 points

En dimension 3 : 10^9 points ! !

Echelles spatiales

Effet local et gradients importants

L'AVC est localisé dans une petite partie du cerveau (au moins au début) et présente des ondes progressives et de forts gradients en espace.

Sait-on résoudre les échelles spatiales ?

Expérimentation numérique : en 1D, résolution du système de l'AVC avec un solveur d'EDO d'ordre élevé => solution « exacte » du système discrétisé en espace

Au moins 1000 points

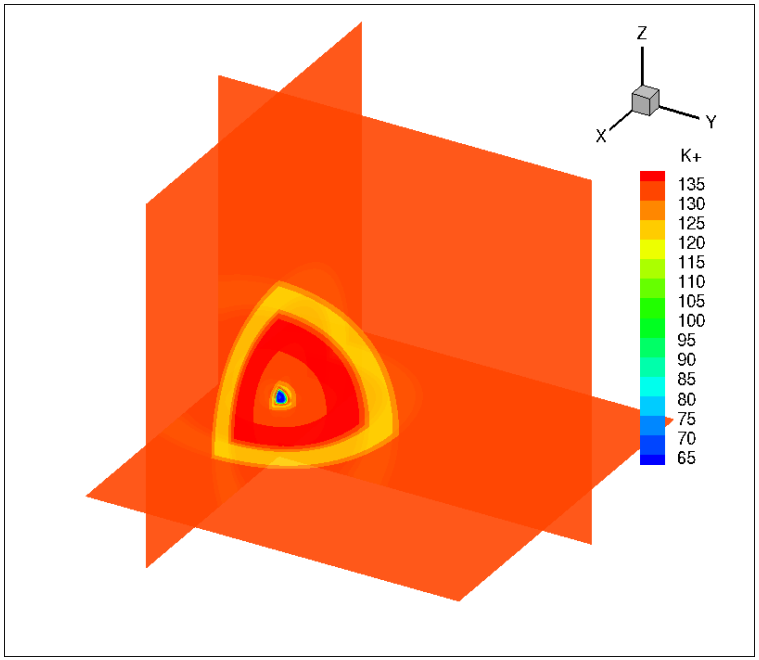
En dimension 3 : 10^9 points ! !

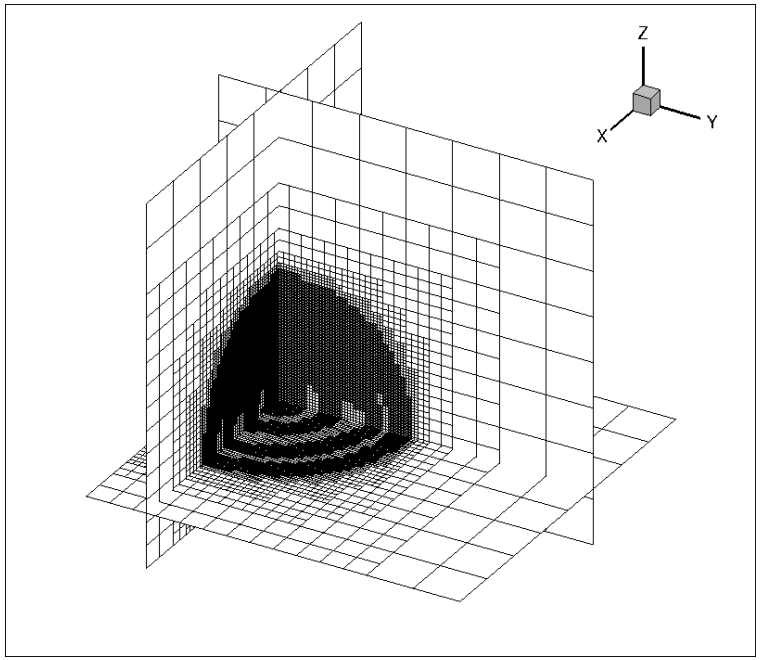
Raffinement de maillage dynamique => Multirésolution

Multirésolution

Intérêts

- Résolution des structures spatiales les plus fines
- Traiter des domaines plus complexes
- Réduire les coûts CPU et mémoire





L'approche multirésolution

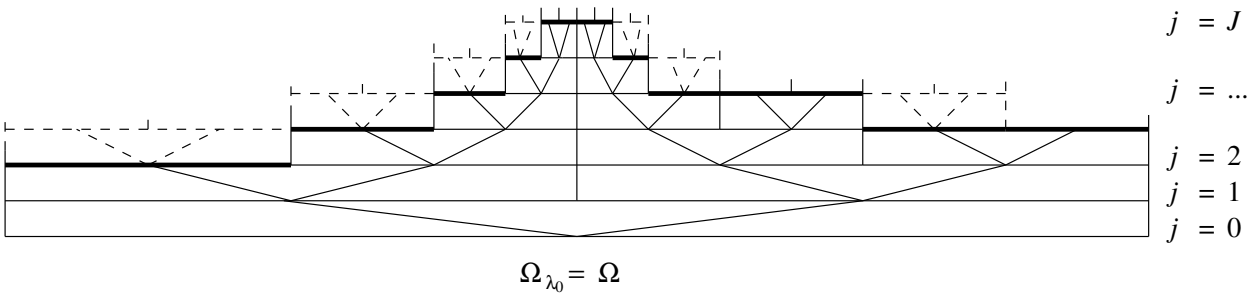
- Fondements théoriques solides basés sur les ondelettes

L'approche multirésolution

- Fondements théoriques solides basés sur les ondelettes

Mode d'emploi

Un ensemble de grilles emboîtées, de la plus grossière à la plus fine



Passage d'un niveau à l'autre :

- Opérateur de projection du niveau j au niveau $j-1$: $R_{j,j-1}$
- Opérateur de prédiction du niveau $j-1$ au niveau j : $P_{j-1,j}$

L'approche multirésolution

- Consistance : $R_{j,j-1} \circ P_{j-1,j} = Id$
- Détails : $d = (Id - P_{j-1,j} \circ R_{j,j-1})u$ où u est défini sur tous les nœuds de l'arbre.
- Valeurs sur une grille = valeurs sur la grille inférieure + détail

L'approche multirésolution

- Consistance : $R_{j,j-1} \circ P_{j-1,j} = Id$
- Détails : $d = (Id - P_{j-1,j} \circ R_{j,j-1})u$ où u est défini sur tous les nœuds de l'arbre.
- Valeurs sur une grille = valeurs sur la grille inférieure + détail

Idée

Raffiner ou déraffiner le maillage en fonction de la valeur des détails.

L'approche multirésolution

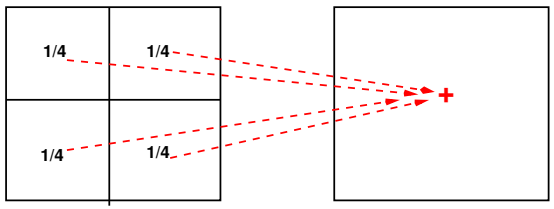
- Consistance : $R_{j,j-1} \circ P_{j-1,j} = Id$
- Détails : $d = (Id - P_{j-1,j} \circ R_{j,j-1})u$ où u est défini sur tous les nœuds de l'arbre.
- Valeurs sur une grille = valeurs sur la grille inférieure + détail

Idée

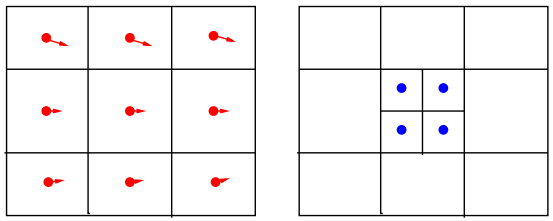
Raffiner ou déraffiner le maillage en fonction de la valeur des détails.

Choix possibles pour $R_{j,j-1}$ et $P_{j-1,j}$:

- $R_{j,j-1}$: moyennes

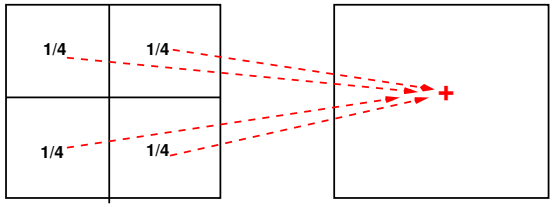


- $P_{j-1,j}$: interpolation quadratique



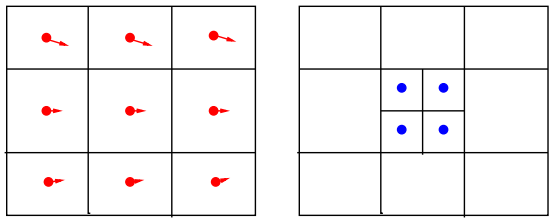
L'approche multirésolution

- $R_{j,j-1}$:



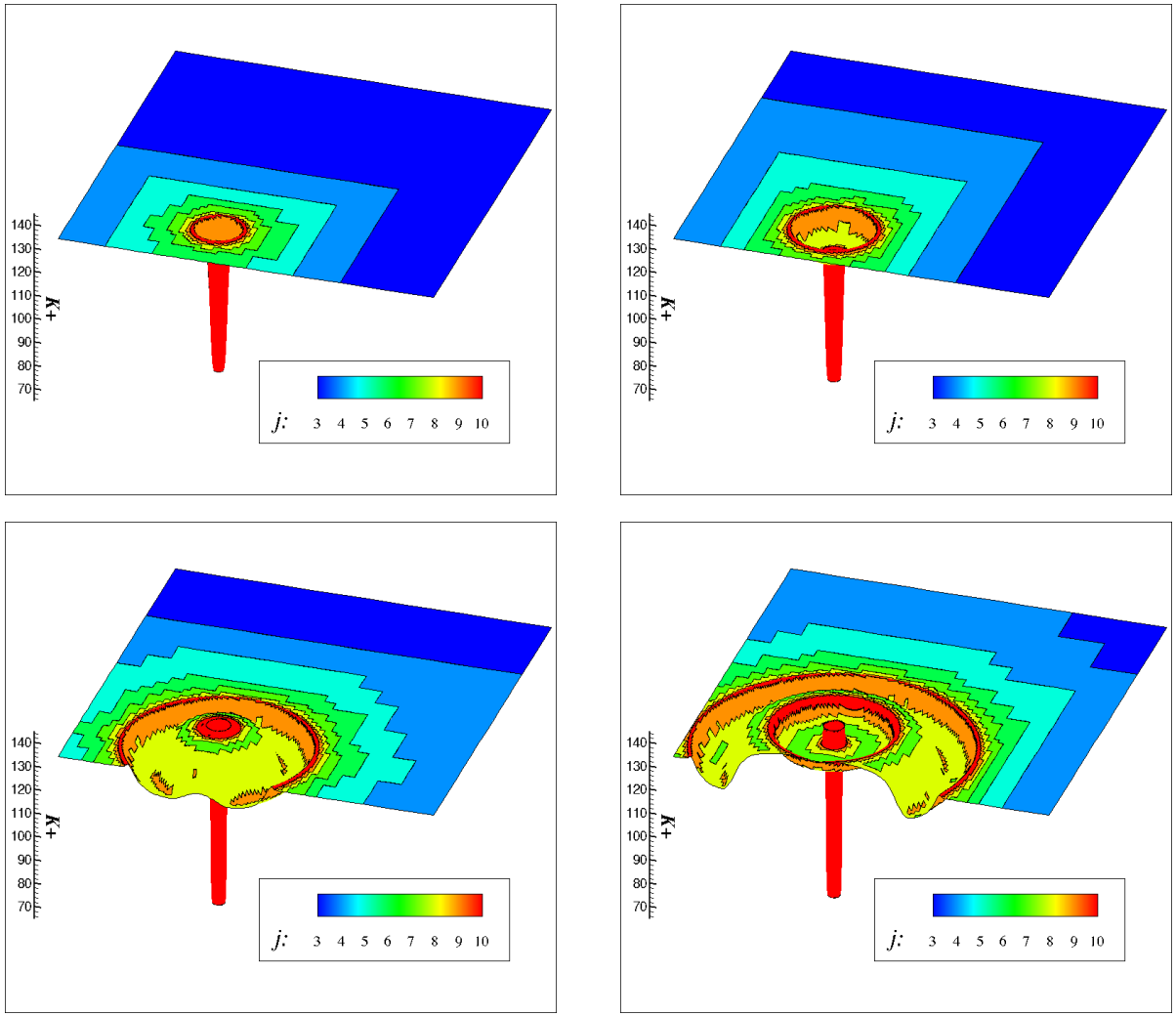
On regarde seulement des nœuds frères.

- $P_{j-1,j}$:



On regarde les cousins et les oncles.

Quelques résultats



$K+$ neuronale à 500, 1000, 2400, 3600 seconds avec 10 niveaux de grille.

Sommaire

1 Contexte et enjeux sociétaux

2 Modélisation de l'AVC

3 Stratégies numériques

4 Détails de la mise en œuvre

5 Problématique de la parallélisation

6 Conclusions et perspectives

Quid de l'implantation ?

=> Méthodes numériques adaptées et efficaces

Quid de l'implantation ?

$=>$ Méthodes numériques adaptées et efficaces

Caractéristiques de la mise en œuvre

- Réaction : opérateur local, parallélisation triviale mais très déséquilibrée
- Diffusion : explicite, presque uniquement des produits matrice-vecteur, simple à paralléliser

Quid de l'implantation ?

=> Méthodes numériques adaptées et efficaces

Caractéristiques de la mise en œuvre

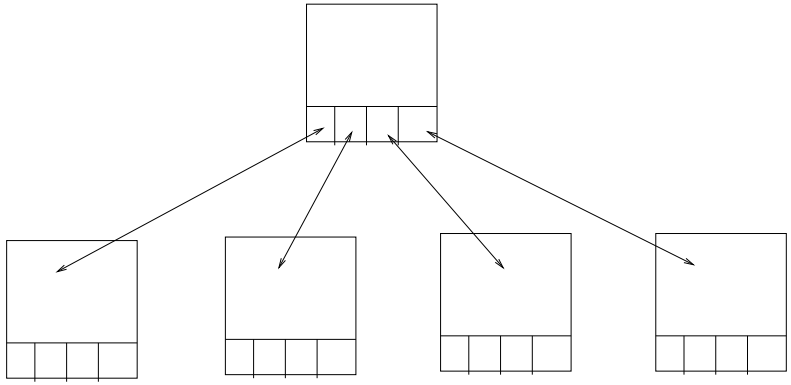
- Réaction : opérateur local, parallélisation triviale mais très déséquilibrée
- Diffusion : explicite, presque uniquement des produits matrice-vecteur, simple à paralléliser

Mais ...

- Les étapes de réaction et de diffusion ont des patterns d'accès aux données complètement différents
- La multirésolution ajoute des problématiques de localité des données très complexes

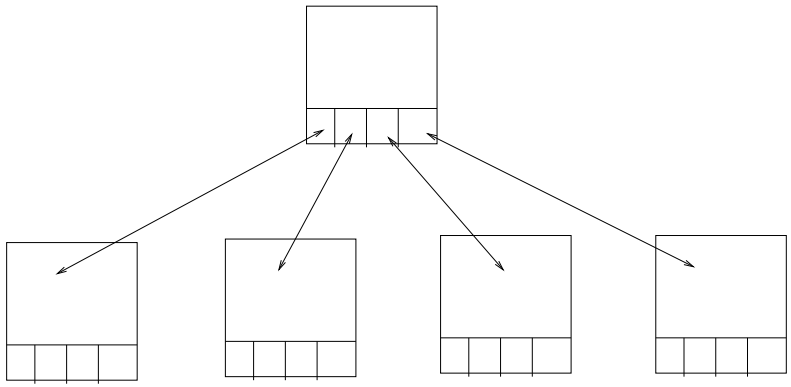
Problématique de la structure de données

La façon classique d'implanter des arbres : utilisation de pointeurs



Problématique de la structure de données

La façon classique d'implanter des arbres : utilisation de pointeurs

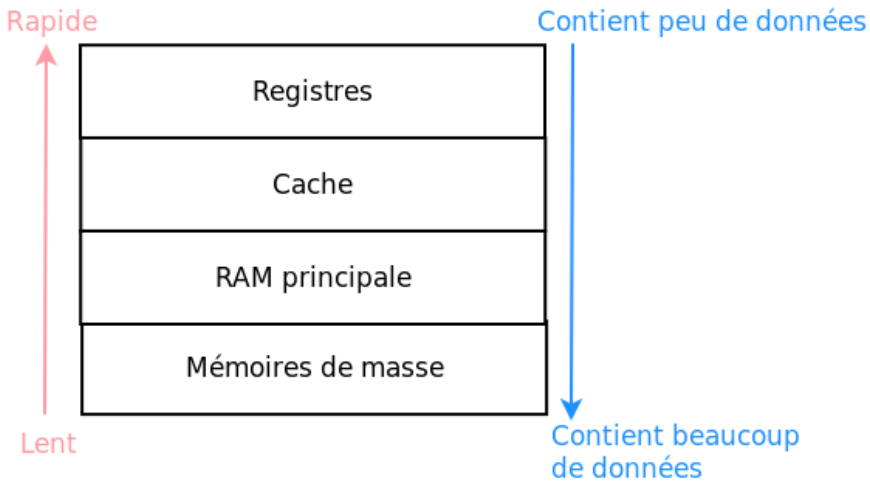


Mais :

- Aucune localité des données,
- Recherche des parents comme les oncles et cousins très lente
- Aucune optimisation possible par les compilateurs

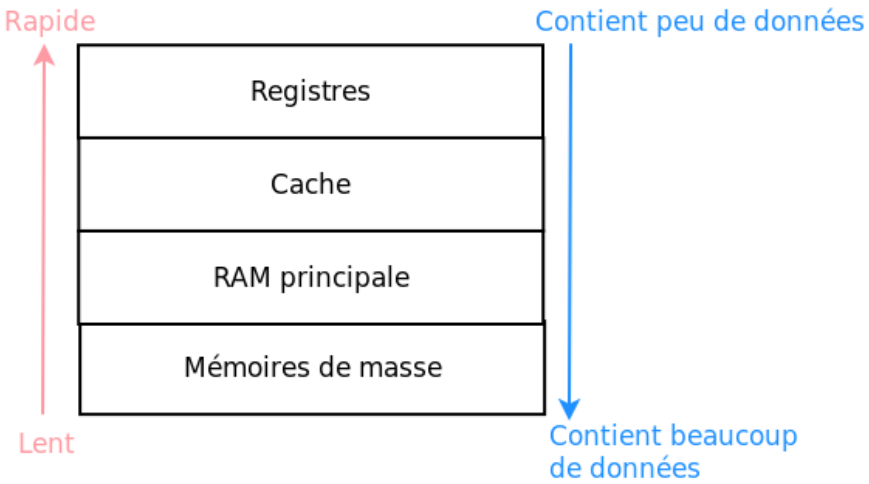
Quelques rappels sur la mémoire

- La mémoire est divisée en RAM puis caches L3-L2-L1 de plus en plus proches du processeur (hiérarchie mémoire)



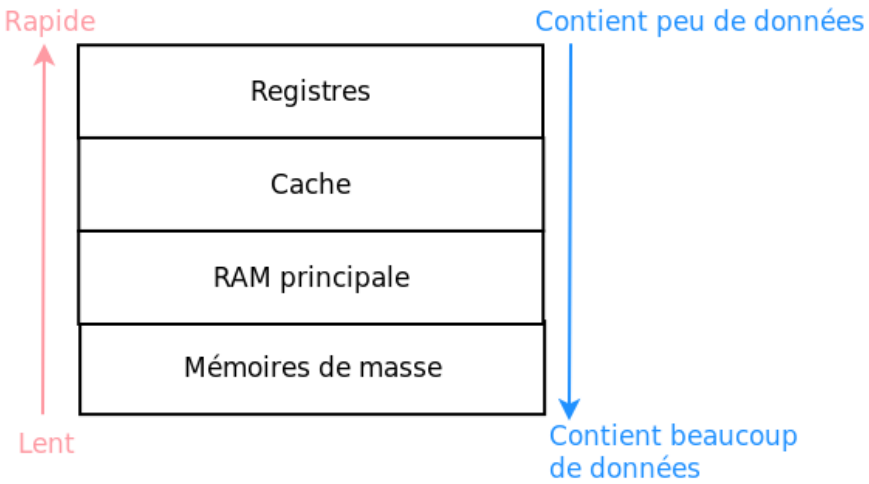
Quelques rappels sur la mémoire

- La mémoire est divisée en RAM puis caches L3-L2-L1 de plus en plus proches du processeur (hiérarchie mémoire)
- Chaque fois que le processeur a besoin d'une donnée, il va la chercher dans les caches
  - soit la donnée est présente dans le cache : cache hit, coût raisonnable
  - soit la donnée doit être récupérée en RAM : cache miss, coût très important



Quelques rappels sur la mémoire

- La mémoire est divisée en RAM puis caches L3-L2-L1 de plus en plus proches du processeur (hiérarchie mémoire)
- Chaque fois que le processeur a besoin d'une donnée, il va la chercher dans les caches
  - soit la donnée est présente dans le cache : cache hit, coût raisonnable
  - soit la donnée doit être récupérée en RAM : cache miss, coût très important
- Les données sont chargées dans les caches par blocs de mémoire consécutive



Localité des données

Localité temporelle

Un programme a tendance à réutiliser les instructions et données qui ont étés accédées dans le passé

```
for (i=0 ; i < 10000 ; i++) {
    s = s + a[i];
}
```

Localité des données

Localité temporelle

Un programme a tendance à réutiliser les instructions et données qui ont étés accédées dans le passé

Localité spatiale

un programme a tendance à utiliser des instructions et des données qui ont des adresses mémoires très proches
un tableau est une structure de donnée respectant le principe de localité spatiale, tandis qu'une liste chainée ou un arbre n'en sont pas

```
for (i=0 ; i < 10000 ; i++) {
    a[i] = a[i] + b[i];
}
```

Localité des données

Localité temporelle

Un programme a tendance à réutiliser les instructions et données qui ont étés accédées dans le passé

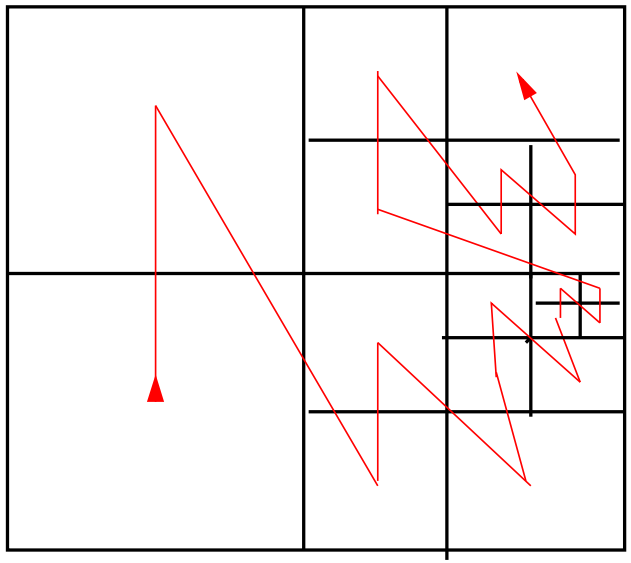
Localité spatiale

un programme a tendance à utiliser des instructions et des données qui ont des adresses mémoires très proches
un tableau est une structure de donnée respectant le principe de localité spatiale, tandis qu'une liste chainée ou un arbre n'en sont pas

Trouver une structure de données qui respecte au maximum ces principes !

Implantation des arbres

Comment stocker des 2-4-8-arbres de manière performante ? Une idée qui vient des bases de données géographiques : un avatar de la courbe de Peano–Hilbert (space filling curve) :



- Respecter le mieux possible la localité en mémoire
- Recherche facile de la parentèle

Sommaire

1 Contexte et enjeux sociétaux

2 Modélisation de l'AVC

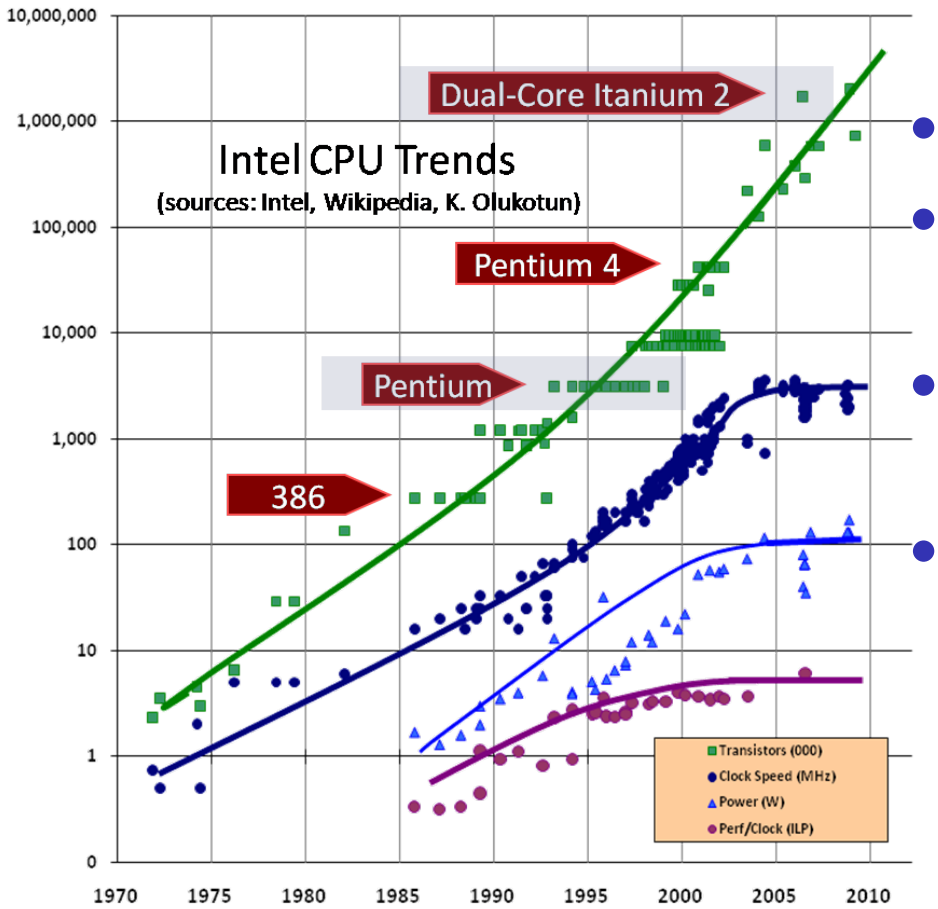
3 Stratégies numériques

4 Détails de la mise en œuvre

5 Problématique de la parallélisation

6 Conclusions et perspectives

Evolutions actuelles

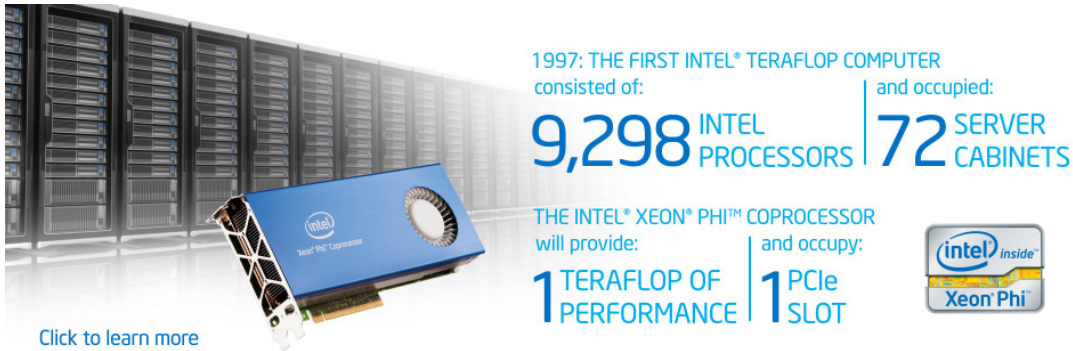


- Fréquence qui stagne voir diminue
- De plus en plus de transistors => de plus en plus de cœurs
- Problèmatique d'accès concurrent à la mémoire
- Architectures hétérogènes : GPU, cartes accélératrices, processeurs ARM ...

Intel Xeon Phi© vs GPU

Many Integrated Core : Xeon Phi© (SE10P/X)

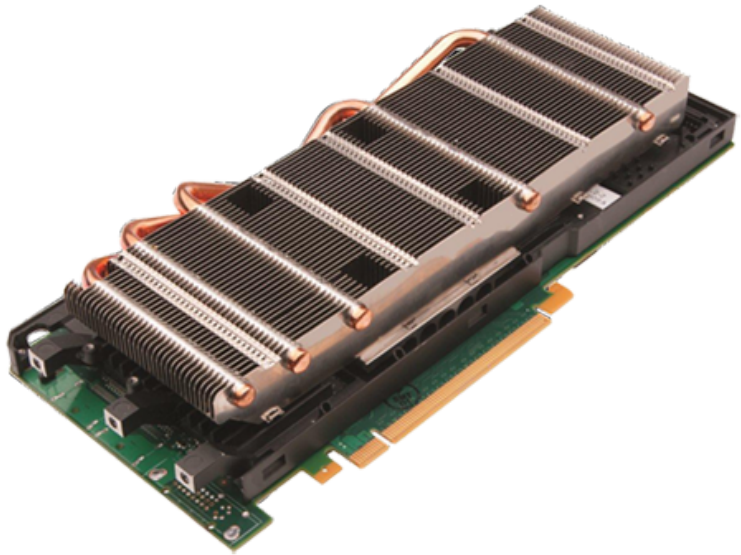
- Co-processeur sur PCIe
- 1.073 TeraFLOP en double précision
- 61 cœurs (1.1 GHz), 8 GB de mémoire (352 Go/s)



Intel Xeon Phi© vs GPU

GPU : Nvidia K20X

- Co-processeur sur PCIe
- 1.317 TeraFLOP en double précision
- 2688 cœurs (735 MHz), 6 GB de mémoire (250 Go/s)



Intel Xeon Phi© vs GPU

Many Integrated Core : Xeon Phi© (SE10P/X)

- Co-processeur sur PCIe
- 1.073 TeraFLOP en double précision
- 61 cœurs (1.1 GHz), 8 GB de mémoire (352 Go/s)

GPU : Nvidia K20X

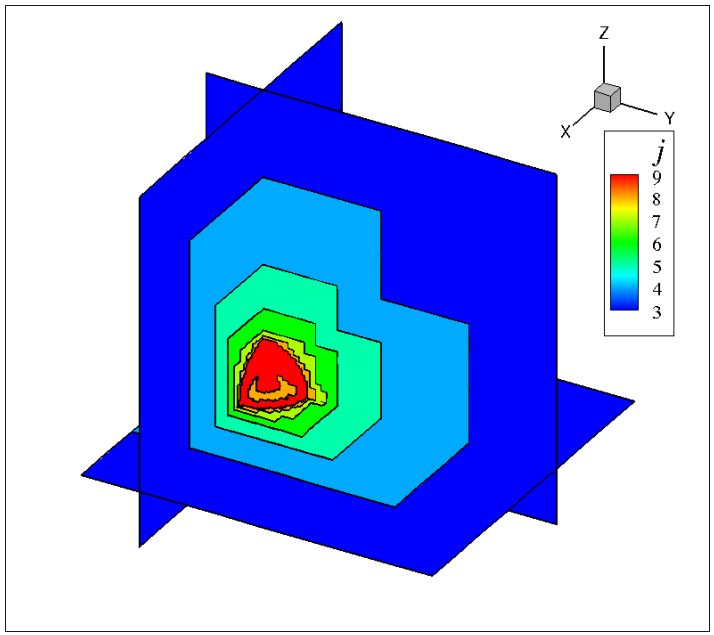
- Co-processeur sur PCIe
- 1.317 TeraFLOP en double précision
- 2688 cœurs (735 MHz), 6 GB de mémoire (250 Go/s)

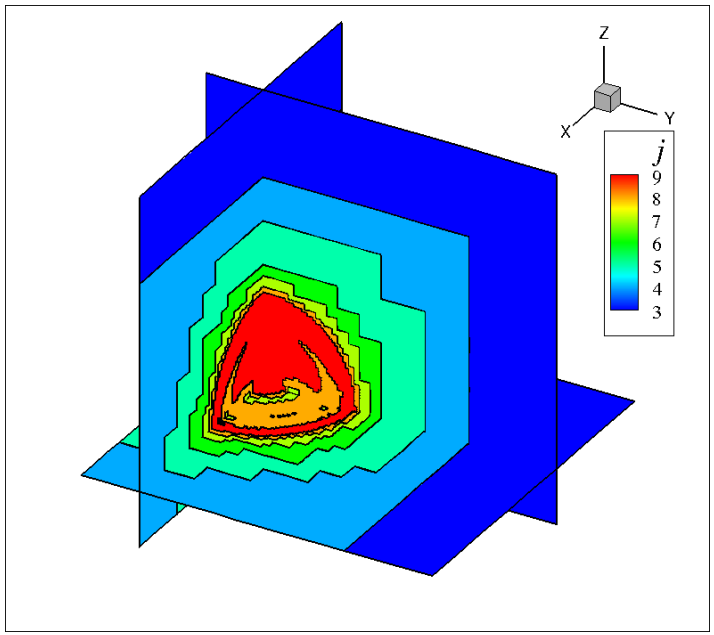
Similarités et différences

- Beaucoup de threads
- Peu de mémoire, coût des transferts mémoire important
- CUDA/openCL pour Nvidia, langages « classiques » pour Intel

Problématique de la multirésolution

- Notion d'arbre (quad-tree, octree) intrinséquement globale
- Application dynamique :
  - le nombre de points/mailles évolue au cours du calcul (augmentation ou diminution)
  - le coût unitaire de la réaction est très variable d'un point à l'autre





Problématique de la multirésolution

- Notion d'arbre (quad-tree, octree) intrinséquement globale
- Application dynamique :
  - le nombre de points/mailles évolue au cours du calcul (augmentation ou diminution)
  - le coût unitaire de la réaction est très variable d'un point à l'autre

=> mémoire partagée dans un premier temps.
Comment gérer l'aspect dynamique ?

Paradigme de programmation par vol de tâches

Principe

Ecrire un programme qui crée des tâches et déclare des dépendances entre elles. Le moteur exécutif décide quelles tâches peuvent être exécutées en parallèle et répartit la charge entre les cœurs.

Paradigme de programmation par vol de tâches

Principe

Ecrire un programme qui crée des tâches et déclare des dépendances entre elles. Le moteur exécutif décide quelles tâches peuvent être exécutées en parallèle et répartit la charge entre les cœurs.

Avantages

- Adapté à l'expression du parallélisme imbriqué : travail sur l'arbre
- Equilibrage automatique de la charge

Paradigme de programmation par vol de tâches

Principe

Ecrire un programme qui crée des tâches et déclare des dépendances entre elles. Le moteur exécutif décide quelles tâches peuvent être exécutées en parallèle et répartit la charge entre les cœurs.

Avantages

- Adapté à l'expression du parallélisme imbriqué : travail sur l'arbre
- Equilibrage automatique de la charge

Langages, bibliothèques

- Cilk `http://cilkplus.org/`
- TBB (Threading Building Blocks) `http://threadingbuildingblocks.org/`
- OpenMP `http://openmp.org`
- Kaapi `http://kaapi.gforge.inria.fr/dokuwiki/doku.php`

Problématique de la vectorisation

Unités vectorielles sur le Xeon Phi

- Retour de la vectorisation comme sur les Cray des années 80.
- Registres et opérations SIMD sur 512 bits : vecteurs de 8 doubles

Exemple de problème

Résolution numérique des EDOs par méthode implicite : calcul de la **matrice Jacobienne** du second membre par approximation numérique :

$$du/dt = F(u)$$

avec

$$J_{i,j} = \frac{\partial F_i}{\partial u_j}$$

On peut, il *faut* vectoriser ce calcul.

Différentes collaborations

Intel et le laboratoire Exascale

- Intel est l'un des 4 partenaires du labo Exascale centré sur les performances des applications sur les futures machines exascale (avec le CEA, l'USVQ et GENCI)
- Optimiser les applications « réelles » dans un environnement exascale.

Différentes collaborations

Intel et le laboratoire Exascale

- Intel est l'un des 4 partenaires du labo Exascale centré sur les performances des applications sur les futures machines exascale (avec le CEA, l'USVQ et GENCI)
- Optimiser les applications « réelles » dans un environnement exascale.

Objectifs à court terme

- Amélioration des algorithmes et de l'implantation pour accéder à des cas plus complexes (maillages plus fins, enrichissement du modèle ...).
- Tests de nouvelles architectures et de nouveaux paradigmes

Différentes collaborations

Intel et le laboratoire Exascale

- Intel est l'un des 4 partenaires du labo Exascale centré sur les performances des applications sur les futures machines exascale (avec le CEA, l'USVQ et GENCI)
- Optimiser les applications « réelles » dans un environnement exascale.

Equipe INRIA MOAIS

- Développeurs de la bibliothèque Kaapi (Kernel for Adaptative, Asynchronous Parallel and Interactive programming)
- Test et comparaison d'algorithmes de vol de tâches

Sommaire

1 Contexte et enjeux sociétaux

2 Modélisation de l'AVC

3 Stratégies numériques

4 Détails de la mise en œuvre

5 Problématique de la parallélisation

6 Conclusions et perspectives

Conclusions et perspectives

Conclusions pour la partie modélisation et numérique

- Développement de stratégies numériques pour la simulation numériques de problèmes multi-échelles en temps et en espace, application aux AVC, mais aussi à la combustion, à la croissance de tumeurs ...
- Simulation numérique 3D d'un AVC dans une géométrie du cerveau réaliste obtenue pour la première fois

Conclusions et perspectives

Conclusions pour la partie modélisation et numérique

- Développement de stratégies numériques pour la simulation numériques de problèmes multi-échelles en temps et en espace, application aux AVC, mais aussi à la combustion, à la croissance de tumeurs ...
- Simulation numérique 3D d'un AVC dans une géométrie du cerveau réaliste obtenue pour la première fois

Perspectives pour la partie modélisation et numérique

Améliorer à la fois les performance numérique et la modélisation :

- Complexifier le modèle : chimiotactisme - conditions aux limites ...
- Paramètres connus de façon très approchée : évaluer la sensibilité du modèle par rapport à ces paramètres

Conclusions et perspectives

Conclusions pour la partie implémentation

- Développement d'un code multirésolution avec TBB, portage sur Xéon Phi
- Optimisation sur de nombreux points : accés mémoire, occupation des threads, vectorisation

Conclusions et perspectives

Conclusions pour la partie implémentation

- Développement d'un code multirésolution avec TBB, portage sur Xéon Phi
- Optimisation sur de nombreux points : accés mémoire, occupation des threads, vectorisation

Perspectives pour la partie implémentation

- Améliorer encore la structure de données (stencil)
- Tests de nouveaux paradigmes : CnC (Intel Concurrent Collections)
- Vers une parallélisation hybride distribuée/partagée

Bibliographie :

[1] Assyr Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6) :2041–2054 (electronic), 2002.

[2] S. Descombes and M. Massot. Operator splitting for nonlinear reaction-diffusion systems with an entropic structure : singular perturbation and order reduction. *Numer. Math.*, 97(4) :667–698, 2004.

[3] S. Descombes and Dumont T. Numerical simulation of a stroke : Computational problems and methodology. *Progress in Biophysics and Molecular Biology*, 97 :40–53, 2008.

[4] Stéphane Descombes, Max Duarte, Thierry Dumont, Violaine Louvet, and Marc Massot. Adaptive time splitting method for multi-scale evolutionary partial differential equations. *Confluentes Math.*, 3(3) :413–443, 2011.

[5] Stéphane Descombes, Thierry Dumont, Violaine Louvet, and Marc Massot. On the local and global errors of splitting approximations of reaction-diffusion equations with high spatial gradients. *Int. J. Comput. Math.*, 84(6) :749–765, 2007.

[6] M-A. Dronne, J-P. Boissel, and E. Grenier. A mathematical model of ion movements in grey matter during a stroke. *J. of Theoritical Biology*, 2006.

[7] Max Duarte. *Méthodes numériques adaptatives pour la simulation de la dynamique de fronts de réaction multi-échelles en temps et en espace*. PhD thesis, Ecole Centrale Paris, 2011.

[8] Max Duarte, Marc Massot, Stéphane Descombes, Christian Tenaud, Thierry Dumont, Violaine Louvet, and Frédérique Laurent. New resolution strategy for multiscale reaction waves using time operator splitting, space adaptive multiresolution, and dedicated high order implicit/explicit time integrators. *SIAM J. Sci. Comput.*, 34(1) :A76–A104, 2012.

[9] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.